Title
The Role of Curvature in Representing Shapes for Recognition

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 15(0)

Author
Kurbat, Matthew A.

Publication Date
1993

Peer reviewed

The Role of Curvature in Representing Shapes for Recognition

Matthew A. Kurbat
Departments of Psychology and Mathematics
University of Michigan, Ann Arbor
330 Packard St.
Ann Arbor, MI 48104
Phone: (313) 764-0318
E-mail: kurbat@cog.psych.lsa.umich.edu

Abstract

Attneave (1954) claimed that approximations made by connecting the points of maximum curvature (MAX points) in a picture were necessary and sufficient for representing shapes for recognition. Lowe (1985) in turn argued that an equally sufficient representation is created by connecting points of minimum curvature (MIN points); hence MAX points are not necessary. However, both Attneave and Lowe neglected the role of curvature concentration in their arguments. It is hypothesized here that for shapes with curvature concentrated at a small number of points, MAX point pictures are far better representations than MIN pictures. More generally, the more curvature was concentrated in fewer points, the greater the advantage of MAX figures over MIN figures in recognizability. This hypothesis was experimentally verified; some implications for shape representation are discussed.

Introduction

Attneave (1954, p.184) argued that information[1] about the shape of an object is concentrated at the points of greatest curvature, or "...at those points on a contour at which its direction changes most rapidly." He supported his argument with two lines of evidence. First, Attneave had subjects approximate outlines of 16 shapes by picking points on which to place dots so as to best approximate the outlines. Subjects generally chose points of greatest curvature (called "MAX points"; see Figure 1a). Second, Attneave connected the thirty-eight points of greatest curvature in a drawing of a sleeping cat; the result is easily recognizable as such (see Figure 1b). Attneave concluded that the information in a drawing consisted of the points of greatest curvature and how they are connected, and that the nervous system stored only this information. Attneave's work has had a major impact on shape representation and recognition research in psychology and computational vision (for a partial review see Quinlan, 1991).

However, Attneave provided no control condition in his original work. Such a condition can be provided by creating a cat picture made by connecting points of minimum curvature[2] ("MIN" points, which occur between each pair of connected MAX points; see Figure 1c). Lowe (1985) created a MIN point cat, and noted that the MAX and MIN drawings appear to be equally recognizable. Thus Lowe concluded that Attneave's original argument was incorrect. Huttenlocher and Ullman (1987) commented on Lowe's observations, and noted that studies such as Attneave's do not address the problem of what information is necessary to recognize an object, but merely demonstrate that certain information is sufficient.

Lowe's argument is incomplete, however, because it neglects the role of curvature concentration in shape representation. What is curvature concentration? Curvature is change of direction of a contour at a given point, so curvature concentration is, intuitively, the extent to which curvature is not equally distributed across all point in a contour. For example, a circle has the lowest possible curvature concentration, because the curvature of a circle is equal at all

[1] Although Attneave originally meant "information" in an information-theoretic sense, it may also be used in a looser, more metaphorical sense in this context (see Quinlan, 1991).

[2] Actually, what I am calling Lowe's MIN point cat was created by connecting the points midway between points of maximum curvature ("MID" points). Strictly speaking, these points need not be points of minimum curvature, but they typically approximate MIN points sufficiently well. In this paper "MIN point" will mean "MID point".

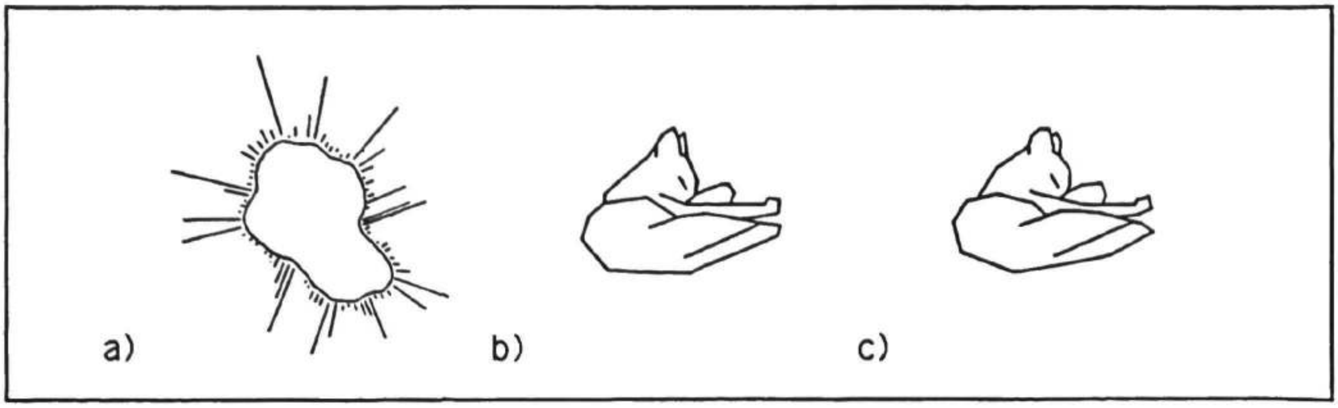

Figure 1: *(a) Subjects approximated the figure shown above with 10 dots. Radiating bars indicate the relative frequency with which various points on the outline were chosen (from Attneave, 1954); (b) drawing made by connecting the points of maximum curvature in a picture of a cat with straight lines (from Attneave, 1954); (c) drawing made by connecting the points of minimum curvature in a picture of a cat with straight lines (from Lowe, 1985; permission for use granted from Kluwer Academic Publishers).*

points. In contrast, a polygon made by connecting a small number of straight line segments has high curvature concentration, because curvature is only found (hence is concentrated) at those points where the straight segments are joined. For example, the curvature of a square is only found at the four corners of the square - the outline of the square changes direction only at those points, and does not change direction in between them [3].

Returning to the question at hand, how does curvature concentration relate to Lowe's argument? If the curvature of a shape is concentrated solely at a few points (i.e., the shape is formed by connecting points with line segments), then a MAX representation using those points recreates the original shape perfectly, while a MIN representation is poor (e.g., see Figure 2). On the other hand, for smoothly curved outlines with many MAX points such as Attneave's cat, curvature concentration is low. In such cases there are at least three reasons why the perceptual difference between MAX and MIN representations should not be nearly so great. First, MAX representations no longer have the advantage over MIN representations of recreating the original shape perfectly - it is impossible to perfectly represent a smoothly curved shape with straight segments. This is another way of saying that smoothly curved objects do not have their curvature concentrated entirely in a finite set of points. Second, if there are many MAX points, then the difference between MAX and MIN representations is not as large as it would be if there were fewer MAX points. More MAX points mean more closely-spaced MAX points, and since MIN points occur between each pair of connected MAX points, the more closely-spaced the MAX points are, the closer the MIN points are to them, and the less difference between MAX point and MIN point representations. Third, if many MAX points chosen are "weak"[4] MAX points, then the difference between MAX and MIN representations is not as large as it would be if those weak MAX points were stronger, since weaker MAX points means less informative MAX points. These last two points are illustrated in Figure 3: in Figure 3a, when only the 26 strongest MAX points are chosen the MAX representation looks much more like a horse than the corresponding MIN representation in Figure 3b; but when 25 weaker MAX points are added, the difference between the resulting MAX and MIN representations (Figures

[3] On a narrower definition of curvature, curvature is undefined at the corners of squares, since the curve is not smooth at such points. However, as discussed further in the discussion section (and Hoffman and Richards, 1984), I am treating non-smooth points like the corners of squares as special cases of curvature maxima.

[4] A "strong" MAX point is one where the curvature of the point is much greater than that of the surrounding contour; a "weak" MAX point is one where the curvature of the point is not much greater than that of the surrounding contour.

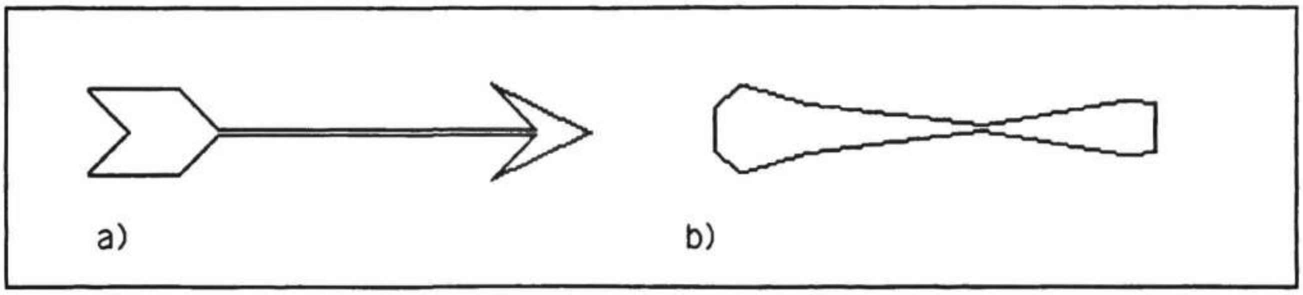

Figure 2: *(a) A computer-generated figure of an arrow approximated by connecting points of maximum curvature; (b) a figure created by connecting midway between points of maximum curvature. In cases where the curvature of a shape is concentrated in a small discrete set of points, a MAX picture approximates the shape perfectly, while a MIN picture is a very poor representation.*

3c and 3d, respectively) is greatly reduced. In general, the more curvature that is concentrated in fewer points for a given shape, the greater the relative advantage of MAX representations using only those points over MIN ones. Thus the purpose of the following experiment was to reexamine the relative utility of MAX and MIN point representations for recognition, but this time to also manipulate the concentration of curvature. The hypothesis was that the more curvature was concentrated in just a few points, the greater the advantage of MAX figures over MIN figures in recognizability.

Biederman (1988) made a relevant methodological point. In commenting on Lowe's work, Biederman noted (and empirically demonstrated) that objects which appear equally identifiable when viewed casually may differ markedly when presented at brief exposures. In general, there may be a number of paths to visual recognition, and real-time classification (at the speeds humans classify) may obey different constraints from those governing classification at slower speeds. As Biederman argues that casually viewing images will not provide much insight to the operation of real-time mechanisms of human vision, it is important (assuming one is interested in real-time vision) to present pictures for only brief exposure durations.

Method

Subjects

Thirty-six introductory psychology students at the University of Michigan participated as part of a course requirement.

Materials

Approximately equal numbers of objects were chosen for each of three concentrations of curvature: concentrated at a small number of points, intermediate, or well-distributed. For shapes with concentrated curvature, all MAX points were used; for shapes with less concentrated curvature, weak MAX points were omitted. In general, more MAX points were chosen for objects with more distributed curvature - because those objects *had* more MAX points, and because more points were needed to recognizably represent such objects. In total, twenty-eight pictured objects were used to create the test stimuli. Two drawings (one MAX and one MIN) were generated by computer from each pictured object for a total of fifty-six test stimuli. Also, eleven practice and three warmup stimuli were used. Pictures were presented on MacII computers.

Procedure and Design

Subjects were told that their task was to judge if names and approximated pictures of objects did or did not match. On each trial the name of the object was presented at the top of the screen, and the picture was presented just beneath it. Subjects learned the task using practice trials, and then went on to the test trials. They began with three non-speeded practice trials, in which the name remained on the screen while the picture was being presented. Subjects were asked to respond either "yes" that the name matched the picture, or "no" that it did not. Subjects were instructed to respond "yes" with the "?/" key and "no" with the "z" key (following Ratcliff and McKoon, 1989),

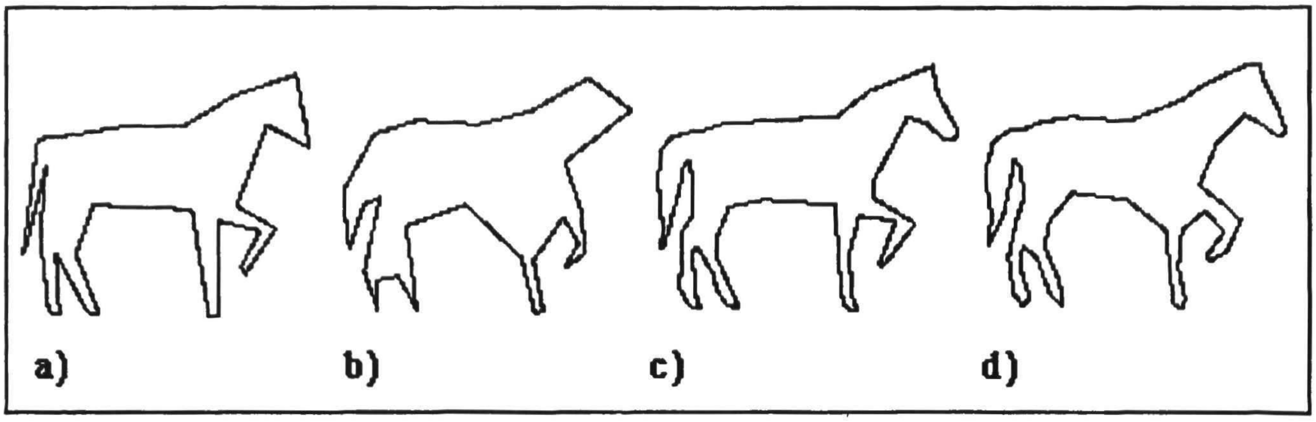

Figure 3: *(a) A computer-generated figure of a horse made by connecting the 26 strongest MAX points; (b) the same horse represented by connecting 26 MIN points; (c) the same horse, now represented by connecting the 51 strongest MAX points; (d) the horse represented by connecting 51 MIN points. When fewer or stronger MAX points are chosen, the advantage of MAX over MIN representations is greater than is it when more or weaker MAX points are added. Thus, the advantage of (a) over (b) is greater than the advantage of (c) over (d).*

using their right and left index fingers respectively. These keys were marked with red dots. Subjects then went on to the eight speeded practice trials (four MAX and four MIN intermixed). For these trials the task was changed so that the name was presented on the screen for only 1.5 seconds; the name was then erased and the picture was presented for 200 msec and then erased. Subjects were again asked to respond "yes" or "no", but were now asked to respond as quickly as possible once the picture appeared on the screen. Finally, subjects went on to three warmup trials and the fifty-six test trials; all warmup and test trials were run consecutively without a break. Procedure for warmup and test trials was identical to that in the speeded practice trials.

The two independent variables were picture type (MAX or MIN) and concentration of curvature (concentrated at a small number of points, intermediate, or well-distributed). All subjects saw all pictures (design was within-subjects). Test trials were presented in two blocks, the order of which was counter-balanced. The order of pictures within each block was randomized. Within each block, half of the pictures were MAX drawings and half were MIN drawings. Each block included the MAX or the MIN picture for each object, but never both. Each picture used in a block was shown twice in its block, once on a true trial and once on a false trial. Also, each name appeared twice in each block, once on a true trial and once on a false trial.

Results

Table 1 shows the mean percent correct and reaction time for "true" trials in each condition. This table shows that there were interactions in the direction predicted between the two independent variables for both reaction time ($F(2,30)=2.36$, $p=.09$) and percent correct ($F(2,30)=45.12$, $p<.0001$). Note that accuracy on MIN pictures for the highest value of curvature concentration was around chance (50%). Also, main effects of picture type were present for both percent correct ($F(1,30)=172.9$, $p<.0001$) and reaction time ($F(1,30)=22.9$, $p<.0001$). Effects of curvature concentration were present for both percent correct and reaction time, and were significant for the former ($F(2,30)=27.5$, $p<.0001$), but not the latter ($F(2,30)=1.1$, $p=.33$). For all these main effects, MAX pictures were either more quickly or more accurately recognized than MIN pictures. However, since the previously mentioned interactions between the two independent variables were present for both reaction time and percent correct, it was impossible to unambiguously interpret the main effects; so Tukey pairwise comparisons between cell means were calculated. By this conservative test, MAX pictures were recognized significantly more accurately than MIN pictures for the

Table 1. Reaction times and percent correct by picture type (MAX or MIN) and concentration of curvature (concentrated at a small number of points, intermediate, or well-distributed - denoted as "3", "2", and "1" respectively).

	Curvature concentration=3	
Picture type	Reaction time (msec)	Percent correct
MAX	657.7	96.8
MIN	938.2	47.9
	Curvature concentration=2	
Picture type	Reaction time (msec)	Percent correct
MAX	658.3	93.1
MIN	867.3	69.9
	Curvature concentration=1	
Picture type	Reaction time (msec)	Percent correct
MAX	689.2	92.0
MIN	763.7	87.8

highest and intermediate levels of curvature concentration ($p<.01$), and MAX were recognized significantly faster than MIN at the highest level of curvature concentration ($p=.05$). MAX pictures were also recognized more accurately than MIN pictures for the lowest level of curvature concentration, and MAX were recognized faster than MIN at the intermediate and lowest levels of curvature concentration, but these differences were not significant. In general, the predicted trend across conditions was found: an advantage of MAX over MIN pictures (in terms of accuracy and reaction time) was found, and this advantage increased as curvature grew more concentrated.

Discussion

The hypothesis tested was that the more curvature was concentrated in just a few points, the greater the advantage of MAX figures over MIN figures in recognizability. This hypothesis was strongly confirmed. Thus it appears that although Attneave's general intuition was correct, his exact formulation and test of it were not. Information *is* concentrated at points where curvature is greatest - curvature maxima are more informative than other points. However, exactly how informative (relative to a MIN picture) MAX points alone are about the curve depends on how much curvature was concentrated in just a few points. Concentration of curvature can in turn be unpacked into several things: (1) the number and spacing of MAX points; (2) how strong the MAX points are; (3) and how much of the total curvature of the figure is concentrated in the MAX points. Attneave's cat alone provides a poor test of the current hypothesis because its curvature is fairly well-distributed.

One prima facie additional problem for this hypothesis is again provided by Lowe (1985), who noted differences between smooth shapes (like that of a cat) on the one hand and the line segment representations made by connecting MAX and MIN points on the other hand. Specifically, he argued that "the ability to introduce tangent discontinuities[5] into a smooth curve - at maxima of curvature or elsewhere - without seriously affecting recognition is actually an indication that local values of curvature need not match predicted values," (p.58). However, this argument is incomplete for at least two reasons. First, Biederman (1988) showed that recognition latencies and error rates were substantially higher for both MAX and MIN cats relative to a normal line drawing of a sleeping cat, when pictures were presented for 100 msec exposure durations. Thus for real-time object recognition, the introduction of tangent discontinuities to smooth curves *does* seriously affect recognition. Second, if a tangent discontinuity is treated as a special case of a curvature maximum (see Hoffman and Richards, 1984), then Lowe's argument is irrelevant - ***absolute*** values of local curvature would not match predicted values, but curvature ***maxima***

[5] A tangent discontinuity is a point where a curve has a "kink" in it, a point where a contour or line is not smooth. For example, the four corners of a square are tangent discontinuities.

would be in the same places. MAX representations can still be more easily recognized than MIN representation even if both are more difficult to recognize than the original picture.

Finally, the results presented here suggest one precise explanation for why the sorts of vertex-based object recognition schemes discussed in the computer vision literature (as reviewed by Boden, 1987) only apply to recognition of simple shapes like polyhedra (e.g., a cube). As polyhedra are composed of planar faces, such shapes have zero surface curvature[6] at all points except those where three or more faces join (a vertex). Thus it is possible to exactly represent polyhedra solely in terms of types and relative locations of their vertices - their points of greatest surface curvature[7]. However, following the arguments made above, for an arbitrarily curved object such a representation is a poorer (or, in many cases, poor) representation of the shape of the object. This explanation also implies that recognition models (Biederman, 1987; Hummel and Biederman, 1992; Lowe, 1985) that rely heavily on the non-accidental property of cotermination - another name for "vertex" - are limited because smoothly curved objects by definition do not have vertices.

Acknowledgements

Thanks to Doug Medin, Mary Lassaline, and John Kennedy for discussions, to Dave Thau for programming help, and to Rob Gelick for help running this experiment. This work was partially supported by a Regents' Fellowship from the University of Michigan. Permission for use of figures from Attneave, 1954, is not needed because material is now in the public domain.

[6] Surface curvature (or Gaussian curvature) is found at a given point by first finding the directions in which the surface curves most and least, and then finding the curvature of the curves in those directions on the surface and taking their product. For example, on the surface of a cylinder the direction in which the surface curves least is the direction parallel to the axis of the cylinder - and the curvature (k1) in that direction is zero; the direction of greatest curvature is perpendicular to the direction of least curvature, and the curvature (k2) in that direction is the reciprocal of the radius of the cylinder. Since k1 is zero and k2 is positive, the surface curvature (k1*k2) is thus zero everywhere on the surface of the cylinder. For more information about surface curvature, see Hoffman and Richards, 1984.

[7] As in the previous paragraph, I am treating vertices - tangent discontinuities - as curvature maxima for polyhedra, since they are the only points on polyhedral surfaces at which both k1 and k2 (see previous footnote) are non-zero.

References

Attneave, F., 1954, "Some informational aspects of visual perception," *Psychological Review*, **61**, 183-193.

Biederman, I, 1987, "Recognition-by-components: A theory of object recognition," *Psychological Review*, **94**, 115-147.

Biederman, I., 1988, "Aspects and extensions of a theory of human image understanding," in Z. Pylyshyn (Ed.), *Computational processes in human vision: An interdisciplinary perspective*, Ablex Publishing Co., Norwood, NJ.

Boden, M., 1987, *Artificial intelligence and natural man*, Basic Books, New York.

Hoffman, D.D., & Richards, W.A., 1984, "Parts of recognition," *Cognition*, **18**, 65-96.

Hummel, J., & Biederman, I., 1992, "Dynamic binding in a neural network for shape recognition," *Psychological review*, **99**, 480-517.

Huttenlocher, D., & Ullman, S., 1987, "Recognizing rigid objects by aligning them with an image," MIT AI Memo 937.

Lowe., D., 1985, *Perceptual organization and visual recognition*, Kluwer Academic Publishers, Boston.

Quinlan, P.T., 1991, "Differing approaches to two-dimensional shape recognition," *Psychological Bulletin*, **109**, 224-241.

Ratcliff, R., & McKoon, G., 1989, "Similarity information versus relational information: differences in the time course of retrieval," *Cognitive Psychology*, **21**, 139-155.